Biomarker Selection for Diagnostic Panel

Background

Based on our discussion with endometriosis researcher Dr. Idahliz Flores, our team knew that the best diagnostic test for endometriosis would utilize a combination of biomarkers from both peripheral blood and menstrual blood samples. However, we needed a way to decide which biomarkers to include in our diagnostic panel. Previous work used classification and regression tree analysis to determine that a four biomarker diagnostic panel, including cancer antigen-125, macrophage chemotactic protein-1, leptin, and macrophage migration inhibitory factor for endometriosis could diagnose 48% of subjects with 93% accuracy (Seeber et al., 2008). Similarly, logistic regression and least squares support vector machines have been used to create a different four biomarker diagnostic panel, including annexin V, vascular endothelial growth factor, cancer antigen-125, and glycodelin, which could diagnose endometriosis with a sensitivity of 81-90% and a specificity of 63-81% (Vodolazkaia et al., 2012). Inspired by this work and the advice of Dr. Flores, our team decided to create a script that determines the best combinations of biomarkers by calculating and combining odd ratios based on biomarker sensitivities and specificities reported in the literature (*Figure 1*).

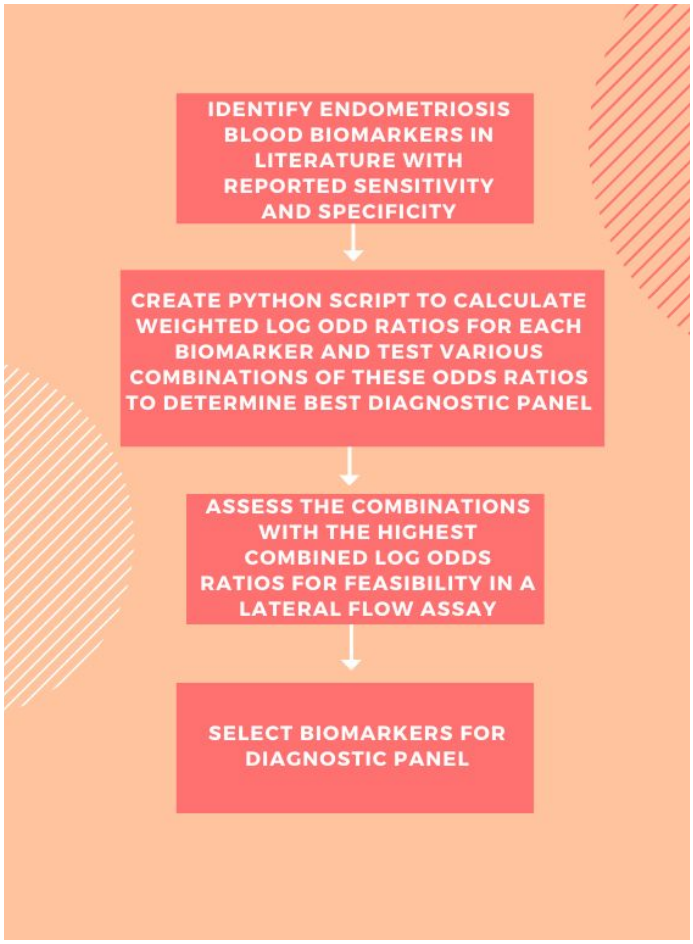


Figure 1: Workflow for Selection of Biomarkers for Diagnostic Panel

Selecting Biomarkers for Analysis

Our approach depends on previously reported biomarker thresholds for endometriosis with their corresponding sensitivities and specificities. Therefore, we conducted a literature search to find biomarker candidates to include in our analysis. We used the following criteria (*Figure 2*):

1. **The study must include individual biomarker thresholds and corresponding sensitivities and specificities.**
   In order to calculate odds ratios, we need the sensitivity and specificity values for the biomarker as a diagnostic for endometriosis. Additionally, we need to know what thresholds were used to calculate these sensitivities and specificities so that if the biomarker is selected for our final panel, we can design the assay to detect that threshold.
2. **The study cannot solely focus on severe to moderate endometriosis.**
   Since our goal is to create a diagnostic for all women with endometriosis, we decided to exclude studies that only included women with severe to moderate endometriosis.
3. **The study must use menstrual effluent and/or peripheral blood samples as sample sources for the biomarkers.**
   Our diagnostic is currently centered around using menstrual effluent to non-invasively detect endometriosis. Therefore, it was important to us that the sample source in the studies is also menstrual effluent so we can apply their thresholds to our design. However, after a discussion with researcher Dr. Flores, we decided to also include peripheral blood samples, as the levels of biomarkers in peripheral blood are likely similar to that in menstrual effluent (Personal Interview, Dr. Idahliz Flores, July 2020) and including peripheral blood increases the number of biomarkers we can include in our analysis.

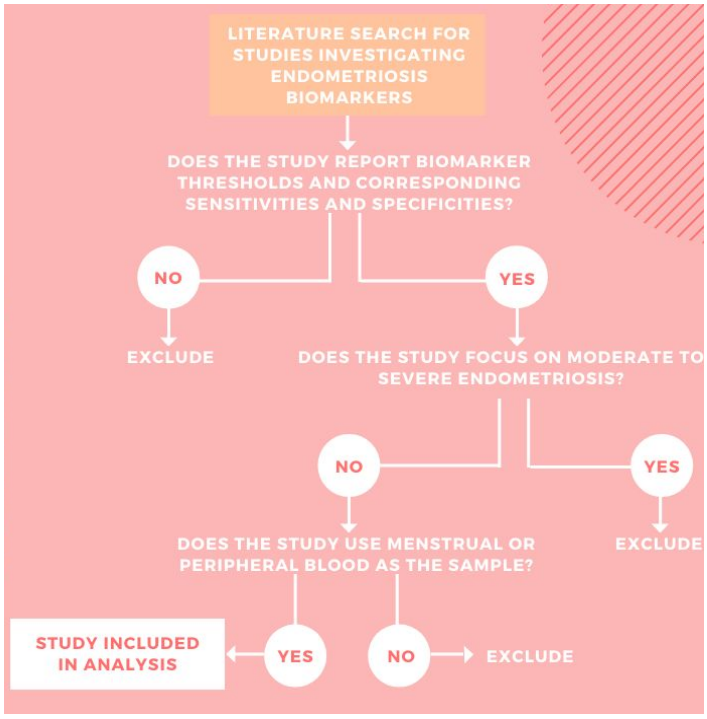


Figure 2: Selection Process for Studies to Include in Analysis

After conducting this review, we identified 12 biomarkers to include in our analysis (*Table 1*).

Biomarker	Sample Size	Endometriosis Sub-sample	Negative Control Sub-sample	Threshold	Specificity	Sensitivity	Source
IL6	91	56	35	2 pg/ml	0.67	0.9	Bedaiwy et al., 2002
IL6	91	56	35	4 pg/ml	0.8	0.85	Bedaiwy et al., 2002
IL6	91	56	35	7.5 pg/ml	0.87	0.8	Bedaiwy et al., 2002
IL6	156	76	80	125 pg/ml	0.95	0.5	Malutan et al., 2015
IL1B	110	56	54	7 pg/ml	0.851	0.751	Malutan et al., 2015
IL8	91	71	21	25 pg/mL	0.714	0.81	Ohata et al., 2008
Glycodelin A	68	48	20	108.5 ng/ml	0.75	0.917	Mosbah, Nabiel, & Khashaba, 2016
TNFa	146	72	74	100 pg/ml	0.729	0.527	Malutan et al., 2015
TNFa	61	30	31	30 pg/ml	0.7742	0.6333	Galo et al., 2005
VEGF	195	103	92	3.88 pg/ml	0.8	0.74	Vodolazkaia et al., 2016
IGFBP1	14	7	7	14.21	0.917	0.875	Gregerson & Metz, unpublished data
CA125	262	161	101	30 units/ml	0.97	0.789	Kitawaki et al., 2005

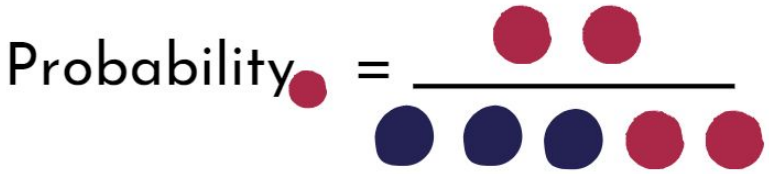
Table 1: **Biomarkers, Threshold Values, Sensitivities, and Specificities Reported in Included Studies**
IL = interleukin, TNF = tumor necrosis factor, VEGF = vascular endothelial growth factor, IGFBP = insulin growth factor binding protein, CA = cancer antigen

Statistical Methods

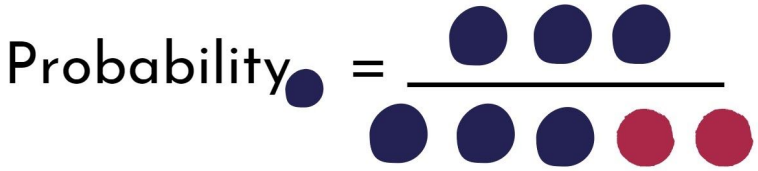
In order to determine the best combination of biomarkers, we considered all possible combinations of three to six biomarkers from the list of 12 biomarkers and identified the combination that provided the highest combined log odds ratio.

Odds ratios are a measure that is often used in clinical diagnostics to determine the association between an exposure and an outcome. In this case, the exposure would be the amount of a biomarker present in an individual's peripheral blood or menstrual effluent, while the outcome would be whether or not that individual has endometriosis. To illustrate this concept, suppose you have the following group of samples:

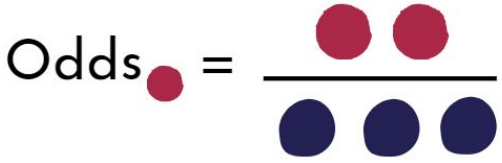
The pink dot represents something happening, such as a woman having endometriosis. The blue dot represents something not happening, such a woman not having endometriosis. The probability of a woman having endometriosis is then:



While the probability of a woman not having endometriosis is:



If we take the ratio of these two probabilities, that is, the probability of a woman having endometriosis divided by the probability of a woman not having endometriosis, we get the *odds* of a woman having endometriosis.



An *odds ratio* is a ratio of odds, a parameter frequently used in diagnostic settings to describe the association between and exposure and an outcome. Odds ratio can be calculated using sensitivity and specificity (*Equation 1, Table 2*).

$$\text{Odds Ratio} = \frac{\text{Sensitivity x Specificity}}{\text{(1-Sensitivity) x (1-Specificity)}}$$

i)

Sensitivity	The proportion of endometriosis subjects that our model correctly identified as having endometriosis
Specificity	The proportion of negative controls that our model correctly identified as healthy

Table 2: Definition of Sensitivity and Specificity

In order to calculate the odds ratio for each biomarker at its reported threshold value, we used the reported sensitivity and specificity. Odds ratio values range from 0 to infinity, where a value of 0 to 1 indicates a negative correlation between the exposure and event, 1 indicates that there is no correlation between the exposure and the event, and 1 to infinity indicates a positive correlation. A negative correlation means that the event is less likely to occur given the exposure while a positive correlation indicates that the event is more likely to occur given the exposure. It is important to note that odds ratios describe associations and not causations.

There is asymmetry in the odds ratio scale, meaning there are more values that indicate a positive correlation (1 to infinity) than values that indicate a negative correlation (0 to 1). This asymmetry makes it difficult to compare odds ratios. Therefore, we took the natural log of the odds ratio (ln(OR)), creating a normal distribution.

Interpreting log odds ratios is notoriously difficult (Osborne, 2018), so we only used this statistic to compare relative performance of biomarker combinations and cannot use this parameter to describe diagnostic accuracy.

In order to determine the ln(OR) of combinations of biomarkers based on the ln(OR) of individual biomarkers in each combination, standard error is used to weight the ln(OR) for each biomarker. This is important because the study sample sizes different between each study included in this analysis. Once the individual ln(OR)s are weighted, the weighted ln(OR) for the biomarkers in each combination are summed and divided by the sum of the weights for each biomarker (*Equations 2,3, 4 & 5*).

$$\text{SE}\,[\ln(\text{OR})] = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

ii)

$$\text{Weight} = \frac{1}{SE^2}$$

iii)

$$\text{Weighted ln(OR)} = weight \times ln(OR)$$

iv)

$$\text{Combined Weighted ln(OR)} = \frac{\sum(weighted\ ln(OR))}{\sum(weights)}$$

v)

a = True Positive Rate
b = False Negative Rate
c = False Positive Rate
d = True Negative Rate

Since calculating all the possible combinations of biomarkers by hand would be time consuming, we created a Python script using numpy, pandas, and itertools to calculate and weight the ln(OR)s for each biomarkers, calculate all possible combined ln(OR)s for combinations of three to six biomarker using a loop, and output the top 15 combined odd ratios.

All statistical methods used were derived from *Epidemiological Research Methods* by Donald R. McNeill, 1996. All calculations were performed using numpy, pandas, and itertools in Python.

Results

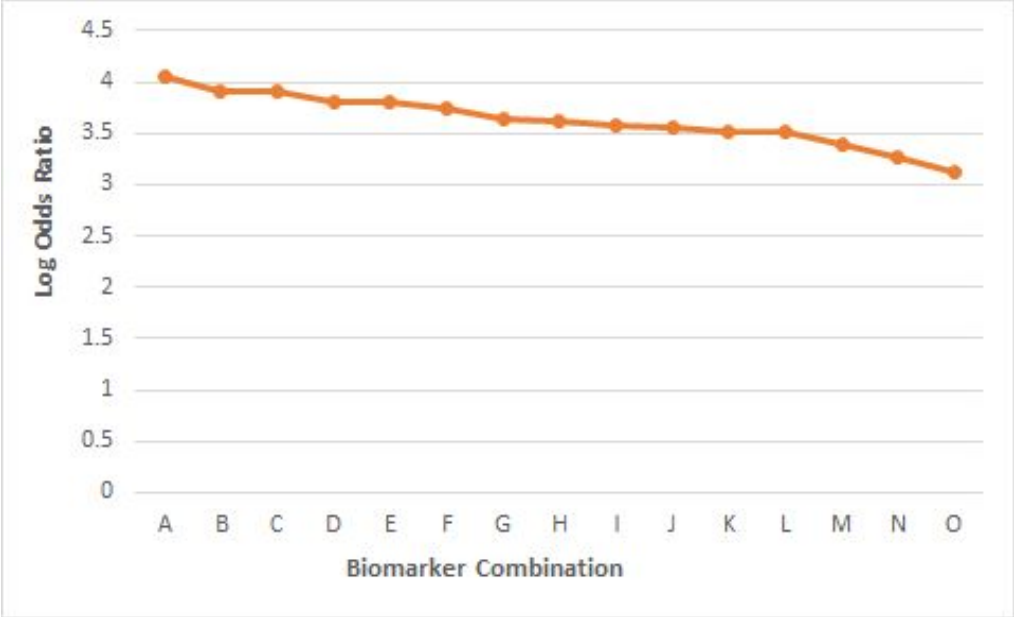
After testing 2,431 combinations of biomarkers, the top 15 combinations were identified and had combined ln(OR)s ranging between 3.130 and 4.045 (*Figure 3A and 3B*). These results indicated that the combination of biomarkers with the highest combined ln(OR) is a three biomarker combination that contains cancer antigen 125 (CA125), insulin growth factor binding protein 1 (IGFBP1), and interleukin 6 (IL6) (*Figure 3B*). Therefore, we decided to include these biomarkers in our diagnostic panel.

A.

Combination	Biomarkers
A	IL6 (7.5), IGFBP1, CA125
B	IL6 (7.5), Glycodelin A, IGFBP1, CA125
C	IL6 (4), IGFBP1, CA125
D	IL6 (4), Glycodelin A, IGFBP1, CA125
E	IL6 (2), IGFBP1, CA125
F	IL6 (2), Glycodelin A, IGFBP1, CA125
G	IL6 (2), IL6 (7.5), IGFBP1, CA125
H	IL6 (2), IL6 (7.5), Glycodelin A, IGFBP1, CA125
I	IL6 (2), IL6 (4), IGFBP1, CA125
J	IL6 (2), IL6 (4), Glycodelin A, IGFBP1, CA125

K	IL6(2), IL6 (4), IL6 (7.5), IGFBP1, CA125
L	IL6(2), IL6 (4), IL6 (7.5), Glycodelin A, IGFBP1, CA125
M	IL6(2), IL6 (4), IL6 (7.5), IL6 (125), Glycodelin A, IGFBP1, CA125
N	IL6(2), IL6 (4), IL6 (7.5), IL6 (125), IL1B, Glycodelin A, CA125
O	IL6(2), IL6 (4), IL6 (7.5), IL6 (125), IL1B, IL8, CA125

B.



C.

Biomarker	Threshold
CA-125	30 units/ml
IGFBP-1	1.60pM
IL-6	7.5 pg/ml
Combined LOR	4.045

Figure 3: Log Odd Ratios Results

(A) Top 15 biomarker combinations. IL6(7.5) = threshold of 7.5 pg/ml, IL(4) = threshold of 4 pg/ml, IL(2) = threshold of 2 pg/ml (B) Log odds ratios of top 15 combinations of biomarkers. (C) Thresholds for best biomarkers in combination with highest log odds ratio.

Creating a Diagnostic Panel for Infertile Women

During our literature search, we found the biomarker prolactin, which was used in one study to diagnose endometriosis specifically in infertile women with a sensitivity of 64% and specificity of 63% at a threshold of 17.5 ng/mL (Mirabi et al., 2019). This biomarker was not included in our main diagnostic panel since it is a biomarker for endometriosis only in infertile women and we aim to use our main diagnostic for both fertile and infertile women. However, we decided to investigate whether adding prolactin to our panel would increase the overall combined ln(OR), which would allow us to create a new diagnostic panel specifically tailored to infertile women. However, the combined ln(OR) decreased upon addition of prolactin to 2.034. Therefore, we decided that **the panel including IL6, IGFBP1, and CA125 was the best combination of biomarkers to diagnose endometriosis in both fertile and infertile individuals.**

Future Directions

This method of analysis relies on publicly available data that was previously collected and analyzed. Ideally, we would have used classification methods on raw data in the form of values for each biomarker in menstrual and peripheral blood samples from women with and without endometriosis. These methods would have enabled us to describe the accuracy, sensitivity, and specificity of our diagnostic panel. Therefore, while our analysis helped guide our decision on which biomarkers to select, this analysis cannot be used to determine the specificity, sensitivity, or diagnostic accuracy of our diagnostic panel.

Future work should focus on collecting menstrual blood samples from women with endometriosis as well as negative controls in order to measure the sensitivity and specificity of using our proposed combination of biomarkers for diagnosing endometriosis.

References

Bedaiwy, M. A., Falcone, T., Sharma, R. K., Goldberg, J. M., Attaran, M., Nelson, D. R., & Agarwal, A. (2002). Prediction of endometriosis with serum and peritoneal fluid markers: a prospective controlled trial. *Human reproduction*, *17*(2), 426-431.

Galo, S., Zubor, P., Szunyogh, N., Kajo, K., Machalekova, K., Biringer, K., & Visnovský, J. (2005). TNF-alpha serum levels in women with endometriosis: prospective clinical study. *Ceská Gynekologie*, *70*(4), 286-290.

Osborne, J. W. (Ed.). (2008). *Best practices in quantitative methods*. Sage.

Kitawaki, J., Ishihara, H., Koshiba, H., Kiyomizu, M., Teramoto, M., Kitaoka, Y., & Honjo, H. (2005). Usefulness and limits of CA-125 in diagnosis of endometriosis without associated ovarian endometriomas. *Human reproduction*, *20*(7), 1999-2003.

Malutan, A. M., Drugan, T., Costin, N., Ciortea, R., Bucuri, C., Rada, M. P., & Mihu, D. (2015). Pro-inflammatory cytokines for evaluation of inflammatory status in endometriosis. *Central-European journal of immunology*, *40*(1), 96.

Martinez, S., Garrido, N., Coperias, J. L., Pardo, F., Desco, J., Garcia-Velasco, J. A., ... & Pellicer, A. (2007). Serum interleukin-6 levels are elevated in women with minimal–mild endometriosis. *Human Reproduction*, *22*(3), 836-842.

McNeil, D. R. (1996). *Epidemiological research methods* (Vol. 322). John Wiley & Sons.

Mirabi, P., Alamolhoda, S. H., Golsorkhtabar Amiri, M., Namdari, M., & Esmaeilzadeh, S. (2019). Prolactin concentration in various stages of endometriosis in infertile women. *JBRA Assisted Reproduction*, *23*(3), 225.

Mosbah, A., Nabiel, Y., & Khashaba, E. (2016). Interleukin-6, intracellular adhesion molecule-1, and glycodelin A levels in serum and peritoneal fluid as biomarkers for endometriosis. *International Journal of Gynecology & Obstetrics*, *134*(3), 247-251.

Ohata, Y., Harada, T., Miyakoda, H., Taniguchi, F., Iwabe, T., & Terakawa, N. (2008). Serum interleukin-8 levels are elevated in patients with ovarian endometrioma. *Fertility and sterility*, *90*(4), 994-999.

Seeber, B., Sammel, M. D., Fan, X., Gerton, G. L., Shaunik, A., Chittams, J., & Barnhart, K. T. (2008). Panel of markers can accurately predict endometriosis in a subset of patients. *Fertility and sterility*, *89*(5), 1073-1081.

Vodolazkaia, A., El-Aalamat, Y., Popovic, D., Mihalyi, A., Bossuyt, X., Kyama, C. M., ... & Meuleman, C. (2012). Evaluation of a panel of 28 biomarkers for the non-invasive diagnosis of endometriosis. *Human reproduction*, *27*(9), 2698-2711.

Vodolazkaia, A., Yesilyurt, B. T., Kyama, C. M., Bokor, A., Schols, D., Huskens, D., ... & Lambrechts, D. (2016). Vascular endothelial growth factor pathway in endometriosis: genetic variants and plasma biomarkers. *Fertility and sterility*, *105*(4), 988-996.